

Languages cooperation across philosophical texts

TheofPhilo, a multilingual thesaurus



Michela Tardella (CNR-ILIESI, Rome)*



DESCRIPTION

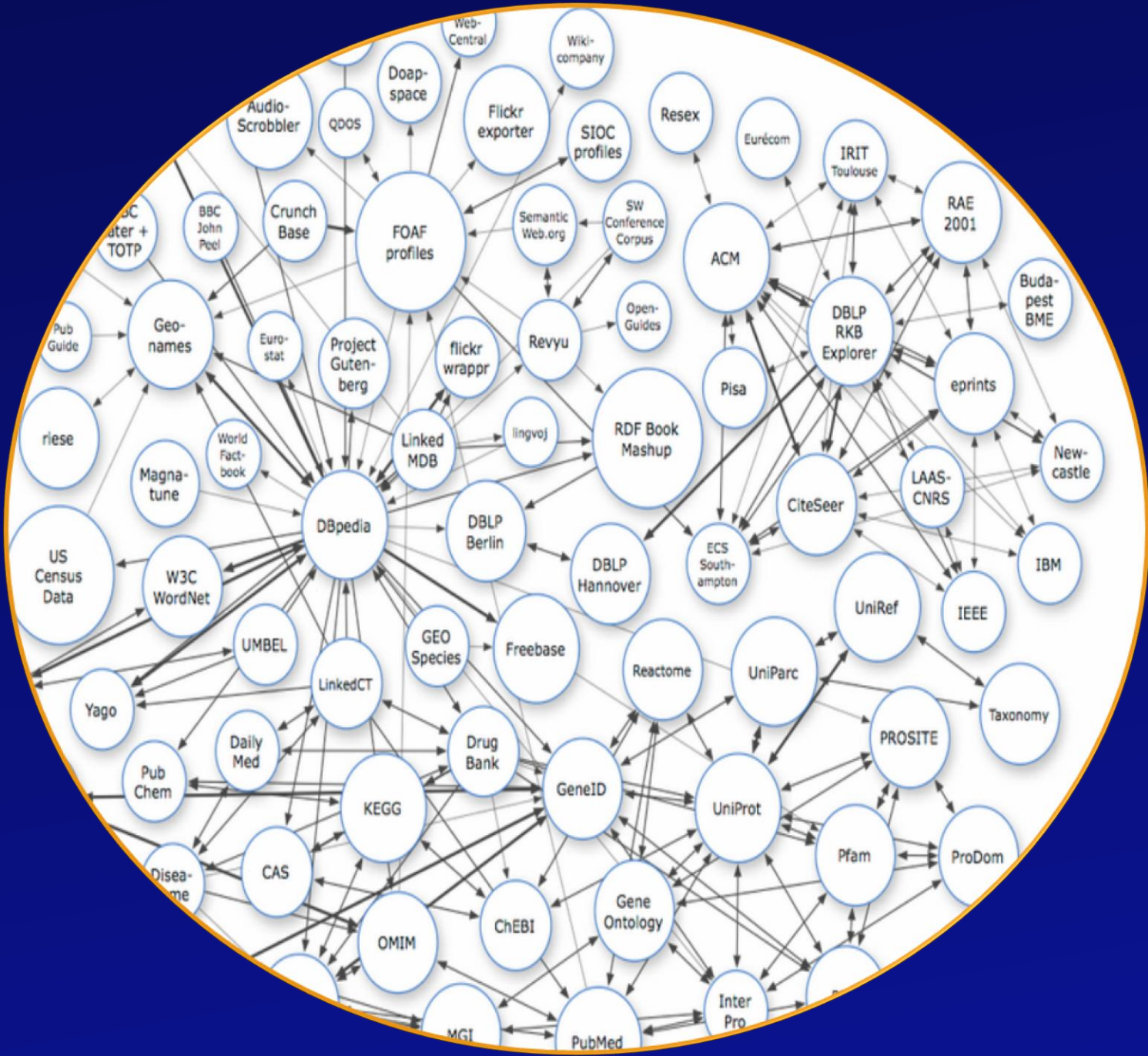
TheofPhilo-
Thesaurus of Philosophy
is the prototype of a digital multilingual thesaurus in the field of philosophy. It is a concept-based oriented collection of terms built on the basis of philosophical texts and dictionaries.

Its purpose is to test the potentialities offered by the interaction between digital tools and infrastructures (i.e. digital archives and libraries) and the tradition of historical and lexicological studies.

It is developed at the *Istituto per il Lessico Intellettuale Europeo e Storia delle Idee* (ILIESI)-CNR characterized by 50 years of expertise in these fields (<http://www.iliesi.cnr.it>).

THE ONTOLOGY

This thesaurus is part of a larger architecture, including a philosophical ontology for the representation of the contents published in the ILIESI' archives. At the present time this ontology is structured according to four categories (*Persons*; *Relevant Concepts*; *Relevant Subjects*; *Sources*) and the related sub-categories. TheofPhilo populates and structures the sub-category of the *Philosophical Subjects*.



PURPOSES

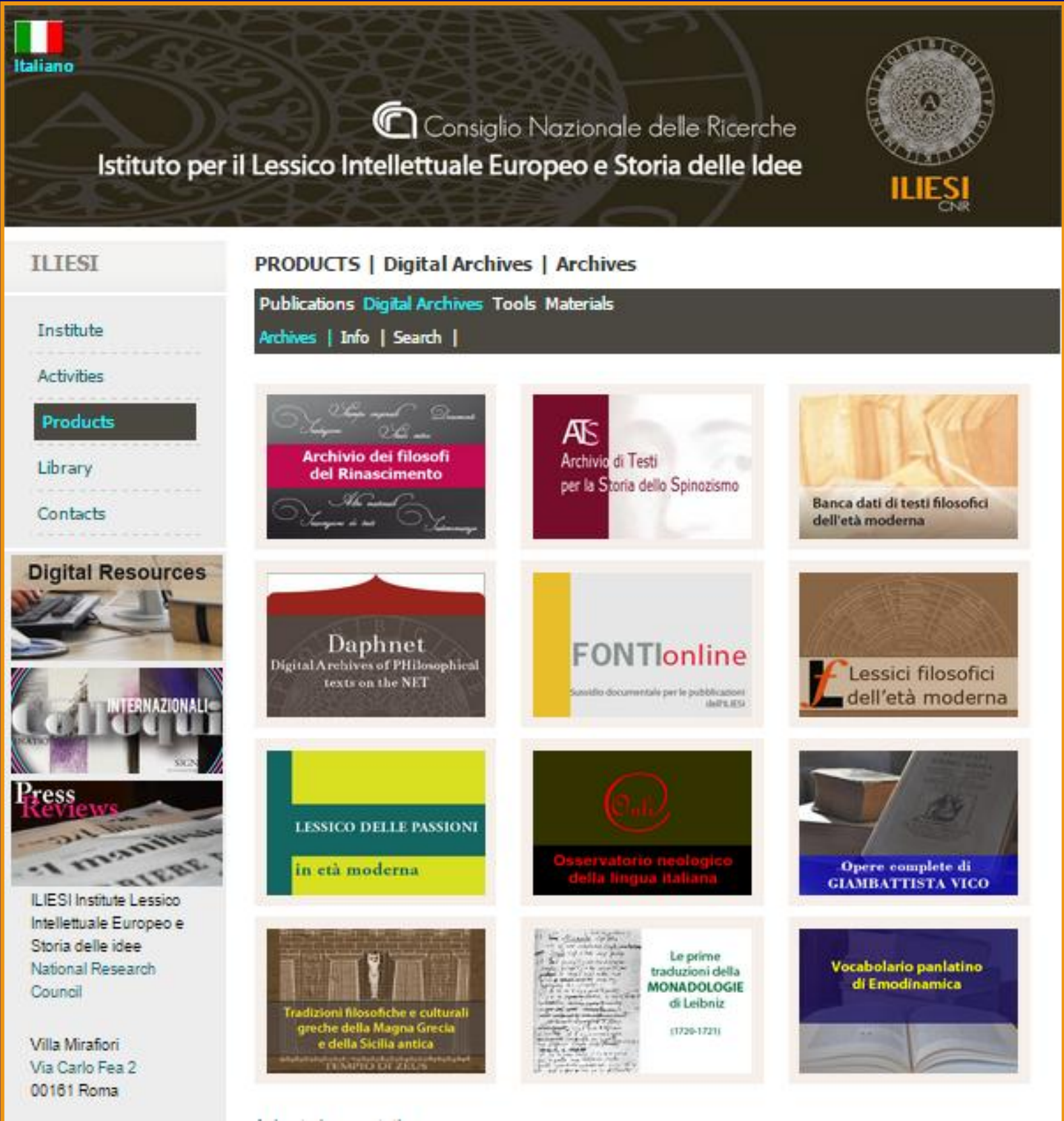
In its first phase of test TheofPhilo has been used for semantic enrichment and Information Retrieval, and has been applied to ILIESI' digital archives, a large collection of open access philosophical texts from ancient Greek philosophy until Enlightenment.

INTERDISCIPLINARITY

The work on TheofPhilo is deeply interdisciplinary. It implies the cooperation of scholars belonging to various disciplines (History, History of Philosophy, Linguistics and Computer Science).

MULTILINGUALISM

The plurality of languages coexisting in ILIESI' digital archives lead us to design a multilingual thesaurus currently consisting of Latin, Greek, Italian, French, and English, but we are planning to include German.



TOOLS 1

For the ontology we use Protégé (<http://protege.stanford.edu>) and the semantic web annotator Pundit (<http://thepund.it>). Pundit is conceived in the increasingly wider context of the semantic web technologies and it allows to produce semantic annotations consisting of a RDF triple. Each triple is a statement made up of a subject (S), a predicate (P) and an object (O). Among the variety of annotations, the most useful to our purposes are those implying triples which connect:

1

A philosophical subject (S) to another philosophical subject (O) by means of the following properties (P):

- ❖ Interlinguistic relations- *Translates_IsTranslationof*
- ❖ Word Families- *Cross-PoS relations*
- ❖ Synonymy- *IsSynonymOf_HasSynonym*
- ❖ Antonymy- *IsAntonymOf_HasAntonym*
- ❖ Homonymy- *IsHomonymOf_HasHomonym*
- ❖ Hyperonymy- *IsHyperonymOf_HasHyperonym*
- ❖ Hyponymy- *IsHyponymOf_HasHyponym*
- ❖ Co-hyponymy- *IsCo-HyponymOf_HasCo-Hyponym*

2

A textual fragment (S) to a philosophical subject (O) by means of the following properties (P):

- ❖ Definition - *Defines_IsDefinedBy*
- ❖ Indirect Definition - *IndirectlyDefines_IsIndirectlyDefinedBy*
- ❖ Extensional Instantiation - *IsAnExtensionalInstanceOf_IsExtensionallyInstantiatedBy*
- ❖ Intensional Instantiation - *IsAnIntensionalInstanceOf_IsIntensionallyInstantiatedBy*
- ❖ Dealings - *DealsWith_IsDealtWithBy*



EXAMPLES

TOOLS 2

We can manage and control the large number of terms by using PHP-MySQL technology.

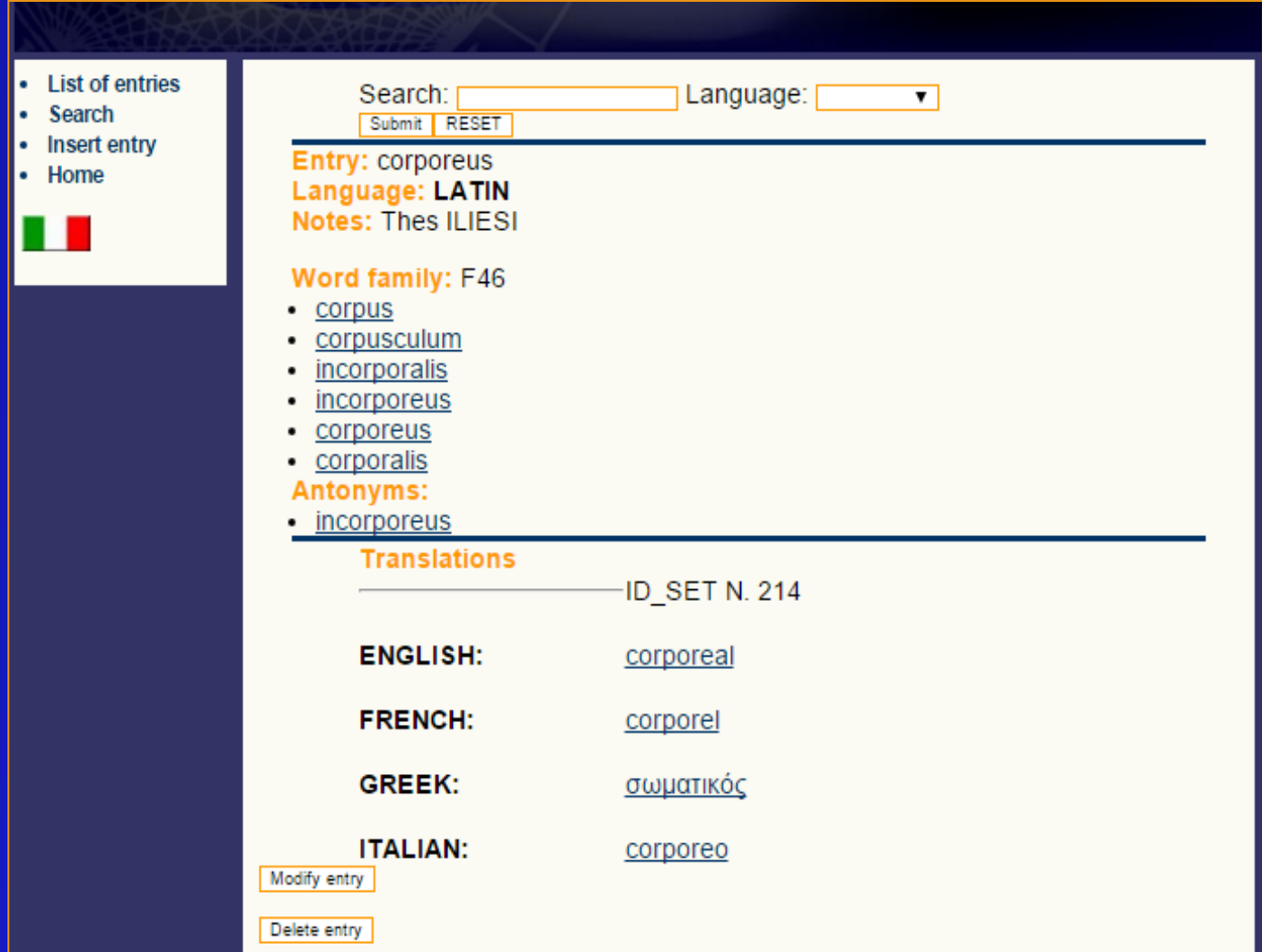


Fig. 1: Query for the Latin term 'corporeus'

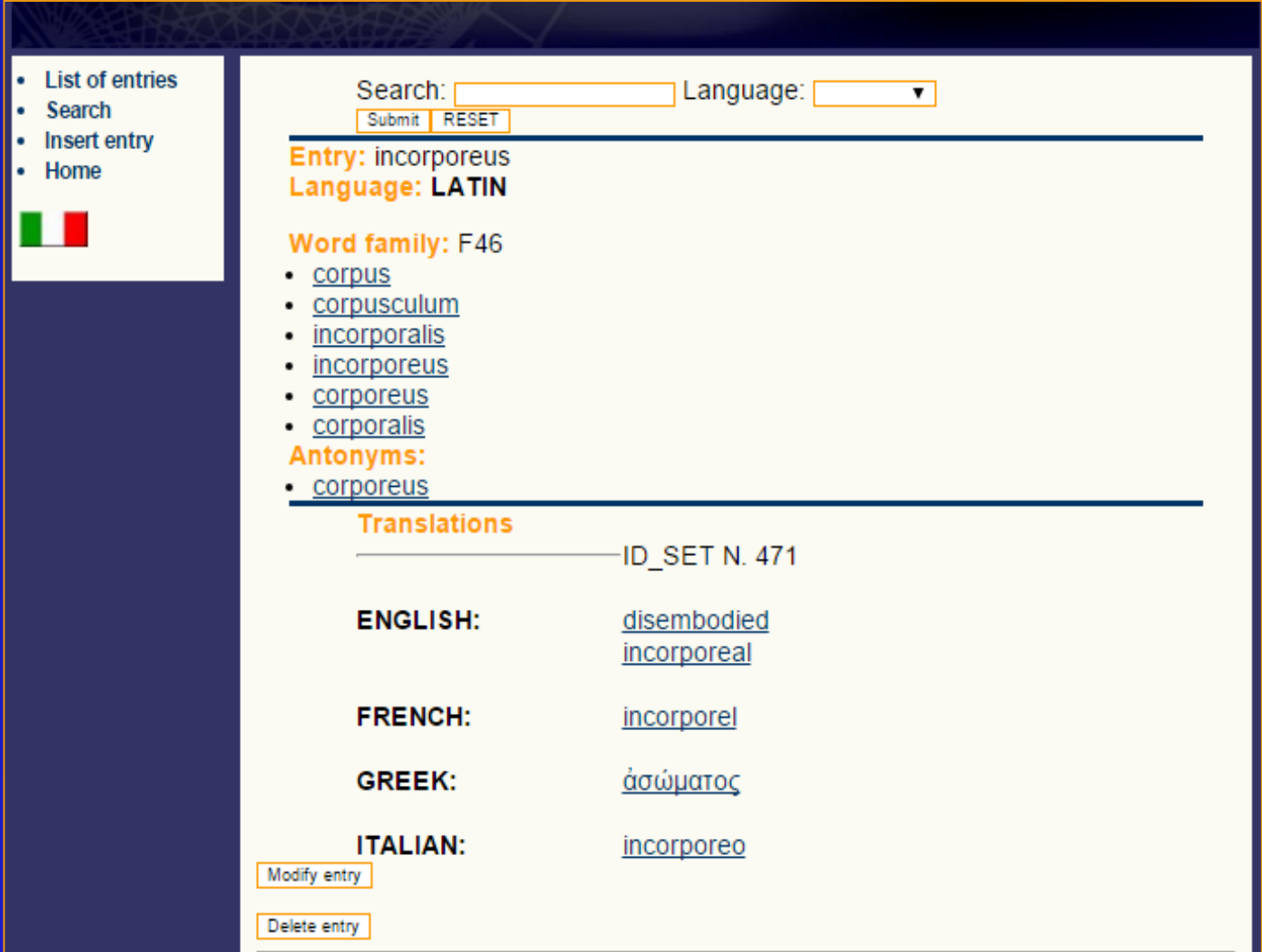


Fig. 2: From 'corporeus' to the antonym 'incorporeus'



Fig. 3: From 'incorporeus' to 'corpus'

METHODOLOGY- Bottom-up model

1. Selection of the meaningful Latin terminology directly from the texts
2. Individuation of word families
3. Selection of the interlinguistic equivalents (Greek, Italian, French and English) by means of multilingual philosophical dictionaries
4. Selection of the intralinguistic relations by means of multilingual philosophical dictionaries
5. Transition from a term-based to a concept-based organization

CRITICAL ISSUES

1. Equivalence and/or translation
2. Heterogeneous approaches of the lexicographical sources
3. Difficulty in managing the equivalences/translations
4. Difficulty in adopting international standards

LIST OF LEXICOGRAPHICAL SOURCES

AAVV, *Enciclopedia filosofica*, Roma 1979.
Abbagnano N., *Dizionario di Filosofia*, Torino 1998.
Bailey A., *Dictionnaire Grec- Français*, Édition revue par L. Séchan et P. Chantraine, Paris 1950.
Baldwin J. M., *Dictionary of Philosophy and Psychology*, Gloucester Mass. 1960.
Cassin B., *Vocabulaire Européen des Philosophies*, Tours 2004.
Forcellini E., *Lexicon totius latinitatis, ab Aegidio Forcellini lucubratum; deinde a Josepho Furlanetto emendatum et auctum; nunc vero curantibus Francisco Corradini et Josepho Perin emendatus et auctius melioremque in formam redactum*, Padova 1940.
Lalande A., *Vocabulaire technique et critique de la philosophie*, Paris 1983.
Liddell G., Scott R., *Greek-English Lexicon*, Rev. by H. S. Jones, Oxford 1968.
Maso S., *Lingua Philosophica Graeca*, Milano-Udine 2010.
Rocchi L., *Vocabolario Greco- Italiano*, Perugia 1993.
Sanesi T., *Vocabolario Italiano-Greco*, Pistoia-Siena 1916.

REFERENCES

Andrews, P. et al. (2011), "A classification of semantic annotation systems", *Semantic Web Journal* 0 (2011): 1-27. <http://www.semanticweb.org/journal/index.php/SemanticWebJournal/article/view/144>
Broughton, V. (2006), "Essential thesaurus construction", London, Facet.
Garshol, L. M. (2004), *Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all*.